Physical geodesy,
an attempt to positioning

Dr. R.F. Rummel

Delft University Press
PHYSICAL GEODESY,
AN ATTEMPT TO POSITIONING

By W. F. Rummel

Delft University Press, 1982
PHYSICAL GEODESY,
AN ATTEMPT TO POSITIONING

Inaugural address
given on the entrance into office
as professor of physical geodesy
at the Department of Geodesy
of the Delft University of Technology
on Wednesday, 20 October 1982

by Dr. R.F. Rummel
Geodesy, as old as science itself, is nowadays almost exclusively viewed as applied science. Almost all of its justification emerges from the needs of a wide community of users, both in science and practice. The applications range from space science, via cadastre to e.g. rural and city planning. Common to all of them is that they circle around the position of a point, the positions of a group of points, or even the shape of the entire surface of the earth. In the sequel it is tried to define in some detail the place physical geodesy takes in this context. The subject of physical geodesy is the study of the structure of the earth's gravity field, an essential contribution to the measurement, determination, and representation of the surface of the earth and its gravity field.

Unfortunately, I cannot save you from the definition of some terms. What is meant by ‘structure of the earth’s gravity field’? Imagine a level and a plumb line, the same sort that is used by the carpenter or the bricklayer, placed at one point. The plumb line defines the local vertical, the level the local horizontal. When repeated in a multitude of points on and outside the surface of the earth, ideally in every point, it will be seen that the levels will form a set of rather smooth closed surfaces, the level or equipotential surfaces. They become increasingly smooth with growing distance from the earth, approaching more and more a spherical shape. The points of suspension of parallel plumb lines, each of them perpendicular to a level surface, form a set of lines, the so-called isocentrical, see Figure 1. The level surfaces and plumb lines can be imagined to be continued inside the earth’s surface too and the only difference is that they will be less smooth.

One of the equipotential surfaces coincides almost perfectly with the surface of the oceans. Deviations due to winds, currents, surges a.s.o. are usually of the order of below 50 cm, which means $10^{-7}$ relative to the earth’s radius. This idealized sea surface can be imagined to be continued beneath the continents and defines the reference level for our heights, the heights above sea level. It is important to note that the definition of heights cannot be given solely geometrically, but has to be embedded into such a physical concept of level surfaces and vertical lines. The question of what is above and beneath can be answered from the direction in which the water flows. It is synonymous to the question of whether a point is located on the same equipotential surface or on one more above or beneath.
We proceed and assign in consecutive order numbers to each of the equipotential surfaces and to each of the isoenethals. This way a coordinated structure is obtained. One speaks of a scalar field. Each point is uniquely defined by a coordinate pair, or a triplet in three dimensions, although by rather complicated curvilinear coordinates. In order to make the characteristics of this field more distinct, we further on assume that the lead weight of the plumb line is not as usual suspended on a rigid cord but instead on an elastic or even better a sensitive spring. We will observe that the spring will be more expanded close to the surface of the earth and slowly reduce to its natural length with increasing distance of it. One could say, a vector field is overlayed to the scalar field. The vectors have varying length, and are, as already mentioned, perpendicular to the equipotential surfaces.

If it is possible to measure, somehow, the structure of the gravity field, expressed for example by the curvilinear coordinates together with the length of the vector at any arbitrary point, isn't it then only logical to try to map them onto another coordinate set more familiar, such as for example three-dimensional Cartesian or ellipsoidal coordinates? If we would succeed to carry out such a mapping or coordinate transformation for the points along the surface of the earth, it would be equivalent to the determination of the figure of the earth, itself. Of course, it would be a hopeless undertaking to try to measure the gravity structure in the entire space outside the earth's surface. A famous theorem by Stokes tells us, however, that once the gravity field is known along one closed surface, e.g., the surface of the earth, or a closed surface in satellite altitude, it can be determined uniquely outside the empty space (if we neglect the mass of the atmosphere). What can be interpreted geometrically as a transformation of curvilinear coordinates, where the metric is derived from the fourth available quantity, the length of the vector, poses physically interpreted a free boundary value problem. In principle one could view this problem statement as a condensed form of the purpose of physical geodesy.

The scalar and vector field structure of the gravity field illustrated by the level, plumb line and spring experiment together with the solution of the geodetic boundary value problem, represents a mathematical model. The parameters describing it have to be estimated from physical observations.

Before I turn to the question of how to actually measure the curved gravity field structure, it seems appropriate to look into its underlying cause. The major cause is gravitation, the mass attraction of the planet earth, with superimposed the tides, due to the attraction of sun, moon, and planets. If we replace, for example, in our model the real earth by a homogeneous sphere or even better by an ellipsoid of the same proportions as the earth, the resulting curved space structure would be much more uniform. A comparison of its equipotential surface at sea level with that of the real earth shows many irregularities and vertical separations of up to 100 m, compare Fig. 2. All these irregularities have to reflect mass inhomogeneities. But where do the latter have their origin? After all, gravitation tries to homogenize irregularities, as all processes that may be comprised within erosion demonstrate every day. One would expect that over the more than four billion years of earth's history the earth would have become nicely smooth and internally symmetrically zoned with the light compounds placed in the outer layers and the most heavy ones in the core. There has to exist a process counteracting gravitation, and as is almost established, it is heat. One could describe the development of the earth's history as a continuous interplay between heat and gravitation acting upon an assemblage of mass particles with characteristic physical and chemical properties. Because of the poor thermal conductivity, the heat in the earth's interior, accumulated in its early history, cannot flow away quickly enough. A geophysical hypothesis, almost generally accepted in its main line, states that hot and heavy material in the mantle expands, becomes lighter and rises, while supported by gravitation, colder material sinks. A convection process, as encountered in various forms in daily life, is set up in the upper mantle. At some places, the ocean ridges, the upward flow even reaches the surface. The outermost part of the earth, the lithosphere, breaks into plates and becomes part of the convection process. The plates, which spread apart in the ridge zones converge again in subduction zones, where the younger and heavier oceanic lithosphere descends into the mantle below the lighter continental lithosphere. It heats up and is absorbed. Results of this interplay heat-gravitation are earthquakes, vulcanism, or large scale continental movements. Results are also mass inhomogeneities inside the earth and the topography, reflected in a complicated anomalous structure of the earth's gravitational...
field. One could say, the gravitational field provides a frozen picture of the process 'heat: gravitation'.

One could now have a bold idea. Stokes' theorem said, that the knowledge of the gravity field along a closed boundary surface suffices to uniquely determine it outside. Is it not also possible to infer the mass distribution inside the earth from the gravity field measured along its boundary and in this way give an almost complete answer to the fundamental questions posed by geophysics? Unfortunately the answer is no, for the second part of the theorem says, there exist infinitely many mass distributions, which possess the same gravity field outside. One is faced with a so-called inverse problem: 'Find the causes from the effects'. The negative answer seems to have almost paralyzed geodesists in the past. I agree, infinitely many possible solutions is a lot, but on the other hand there naturally exist many more mass distributions that do not reproduce the actual gravity field of the earth. Although geodesists are used to work with infinitely dimensional solution spaces in linear adjustment problems, one seems in this context to be chased away by the complexity of the underlying geophysical problem. Geodesists could do more than solely providing a precise approximation of the gravity field to geophysicists. Through the derivation of reliable error measures and their intersection with those to be derived from other geophysical constraints, they could almost certainly contribute to an improved quantification of the problems posed in solid earth physics.

Since geodetic measurements are almost exclusively carried out on or from the earth's surface, an additional contribution to gravity, besides gravitation, comes from the earth's rotation. But even in a 10-5 model set-up this centrifugal part can be modelled rather well. However, the rotational behaviour offers, in addition, the possibility to replace the arbitrary scalar numbers of the isozenithals by a more familiar pair of coordinates, the astronomical co-latitude and longitude, which relate the local vertical to the rotation axis, and the Greenwich meridian. Unfortunately this type of coordinate system choice, related to the rotation axis, meets difficulties with rather fundamental consequences. The geodetic use of any external reference, whether star or satellite, requires the precise knowledge of the variable orientation of the earth in space. Under the combined attraction of sun, moon and planets the rotating and oblate earth performs the famous gyroscopic motion. In addition, an elastic earth with fluid outer core, with oceans and atmosphere and complicated processes going on inside, does not react to these tidal influences in the same way as it would were it a rigid body. For convenience one usually splits the orientation of the earth relative to space into three parts: First, the precession and nutation, the motion of the rotation axis as seen from an observer at rest in space, second, the polar motion, which is the varying orientation of the rotation axis as seen from an observer at rest on earth, and, finally, the varying spin rate. One consequence, for instance of the polar motion, is that latitudes and longitudes are subject to continuous
changes, a prospect not very attractive to the user. To escape these problems, an in principle arbitrary conventional choice of a coordinate axis is defined as a coordinate and similarly a conventional spin rate. Let us assume that the dynamics of the sun, earth, moon system -- that the transformations due to precession-nutation, polar motion, and variations in spin-rate -- are consistently modelled for an oblate and elastic earth model. Then the discrepancies between model and real star observations, or between the model orbit of a satellite and the real satellite tracking data not only help to improve the orientation parameters, but furnish at the same time important information for investigating phenomena such as the elastic response of the earth, core-mantle coupling, or redistribution of mass in atmosphere, oceans, and ground water.

Plumb line, spring, and level were not only introduced to illustrate the complicated structure of the gravity field, they play in fact the same role in the real measurement practice. Most geodetic instruments are oriented along the local vertical with level and plumb line, ultra-sensitive spring gravimeters are employed to measure the variations in the length of the gravity vectors, and together with the level instruments they serve for the determination of the scalar potential differences between the level surfaces, from which afterwards geodetic heights are derived. The length or magnitude of the gravity vector is determined by an improved version of Newton's falling apple, which is said to have made him meditate about the causes of gravitation. Modern devices approach a relative accuracy of 10^{-9}. In essence, drop interval and time of a falling corner-cube reflector in a vacuum chamber is simultaneously measured interferometrically. In a recent development the falling reflector is put into an also free falling vacuum chamber, to eliminate residual air drag in the never perfect vacuum.

This experiment is mentioned in some detail, because it resembles striking analogies with recent developments around another free-fall experiment, that contains enormous prospects for the future: the free fall of a satellite in its orbit around the earth. Up to now only the longwavelength features of the gravity field with extensions of several thousands kilometers could be obtained from the analysis of satellite orbits. Details are difficult to resolve for, first, the details of the gravity field structure are damped out at high altitudes and, secondly, unmodelled non-gravitational forces, such as air drag and solar pressure, restrict the precision. In recent time one focusses on a dramatic increase of the sensitivity. A first requirement is to choose the orbits as low as possible, unfortunately at the expense of much more air drag. The sensitivity for the details of the earth's gravity field is envisaged to be further increased by measuring the relative velocity between two satellites in essentially the same orbit and close together, the so-called satellite-to-satellite tracking. Non-gravitational disturbances can be eliminated by placing each satellite -- as in the case of the free fall apparatus -- in a drag-free shield, a satellite within a satellite with monitored relative motion between the shields. Finally, one could decrease the distance between the two satellites or even place them into one drag-free shield, measuring the differential force acting between the two masses. This is the so-called gradiometer principle. The most recent concept, under consideration, using a test mass in a superconducting environment, could theoretically obtain a sensitivity of 10^{-13} or sense changes of 10^{-13} in g over a distance of 10 cm. Geometrically interpreted such a system directly measures the curvatures of the local curved gravity field structure.

Much was said up to now about the mathematical model. Keywords were curvilinear coordinates and the gravity field, the role of the Stokes theorem, and the determination of the earth's surface interpreted either geometrically as a transformation problem of curvilinear coordinates, or physically, as a free boundary value problem. We also summarized the most prominent measurement methods. One could now deal with the estimation of the parameters describing the shape of the earth and the gravity field from the measurements. In principle, this problem has no solution. For, as expressed by Stokes' theorem, the gravity field structure has to be known at least continuously over one closed surface enclosing all masses, e.g. the surface of the earth. Thus, even all present day available measurements will not suffice. Even if the shape of the earth and the gravity field was modelled not as a continuum but by a finite parameter set, and adequate for a 10^{-8} model, the problem remains drastically underdetermined. Some years ago, the well-known English geodesist Reg and Sprung developed a model, that not only allows to combine all geodetic observables in a unified way, the concept of integrated geodesy, but that at the same time treats the underdeterminancy in an optimal manner. The optimal estimate is attained by singling out among all possible solutions for the gravity field approximation that with minimum norm, so-to-say the smoothest. What was, however, overlooked or at least not emphasized in the past, is the fact that a solution of this type is biased, and this can have enormous consequences for the judgement of the quality of the estimate. As long as the data material is local and sparse, as in the case of terrestrial measurements, the bias shall certainly exceed the envisaged 10^{-8} relative precision. One, and most probably, the only way out of this dilemma is to employ the described highly accurate satellite gravity sensors, to complement the available territorial measurements. Although in satellite altitude, they provide an almost continuous coverage on a closed surface with almost any desired sampling density.

Intuitively one feels that the inclusion of these satellite gravity sensors, above the terrestrial measurements must lead to an accurate estimation of surface and gravity field parameters. But that considerable obstacles have to be overcome, will be sketched by two examples.

When one talks about the realization of a 10^{-8} relative precision concept, which means for instance the determination of the earth's surface accurate to 5 cm in radial direction, then it is mainly the small local details of the gravity field structure, that are needed. The general, global picture
is known rather well. But especially the fine details are considerably damped with increasing distance from the earth, or in other words they become less visible to satellite will sensors. If highly advanced technology nevertheless allows to measure them, another problem is encountered. The model will never perfectly describe the physical observations. There remains always a certain noise. When the details of the gravity field structure on earth are deduced from that observations in satellite altitude, not only the damped gravity signal is amplified, but also the noise contained in the measurements. This results in instabilities or in some cases even in a complete divergence of the processing. Such a phenomenon is typical to problems that were classified as improperly or ill posed by the famous French mathematician Hadamard. Unfortunately they are rather common in natural science.

Whereas strategies are developed in geodesy to treat the instabilities a second complication seems at the moment even more severe. The problem could be described as 'how to make adequate use of the physical information contained in large sets of geodetic observations'. Every precise satellite gravity mission will -- and this is after all the intention -- accumulate hundreds of thousands of observations in a few months. Was it before that I pointed out integrated geodesy is going to produce estimates too biased, when applied to the local and terrestrial measurements, so is it now that integrated geodesy, although ideally in theory can never be applied to such large data sets. For there is no way to solve the equally large system of linear equations. The solution is sought in replacing the discrete estimator by its global continuous counterpart, since the latter permits an analytical solution. What a strange paradox when we remember that in the first step the continuous problem, the geodetic boundary value problem, was approximated by the discrete measurements.

Latest at this point, the stochastic and functional model set-up, presented here, deserves to meet with what is internationally known as 'Delft tradition'. Over the last thirty years, based upon the work of Tienstra, Baarda and his group developed a closed theory in the field of point positioning in two and three dimensions. Functional and stochastic model form a consistent unit. In recent years the theory has been extended as to include space and gravimetric methods, too. However, whereas I sketched the central topics of physical geodesy as ill posed, unstable, and even biased, this theory is properly and stable posed and from the adjustment point of view unbiased. Why not simply apply then the latter? The answer would become rather complex. One might argue for example, that the latter theory is somewhat idealistic from the observational point of view, the former thoughts, starting from the deficiencies on the experimental side, might be too pessimistic. The signs are good for an alliance, the result could be a far reaching model accurate to $10^{-7}$, and includes latest observational strategies.

Although in The Netherlands, the path I have chosen today became rather rugged and stony. As a result almost all my time is used up and there was not sufficiently time to place physical geodesy in the context of its applications, except of what was said about the definition of height systems, and the interrelations to earth sciences. But on the other hand, it seems justifiable that more emphasis was given to the theoretical or model aspects. Geodetic results, if they should be reliable, have to be based upon consistent and adequate models. Only then problems in practice can be treated competently such as:

- the monitoring of vertical crustal movements, caused by groundwater variations, mining or geologic activity, the question of secular sea level changes or the geodetic height datum problems, or

- the geodetic contribution to oceanography: The determination of the equipotential surface at sea level from a satellite gravity mission together with sea gravimetry, and the determination of the actual sea surface from satellite altimetry, where we have already several years of experience should allow to deduce the separation of the two, the sea surface topography to within 10 cm. From it the surface pressure gradients and further the sea surface velocity field can be obtained, which is essential for the study of ocean circulation. Together with the density field it would even allow to study the internal processes such as the transfer of momentum, of heat or of chemicals, subjects of enormous scientific and practical relevance.

- or, still in the beginning, the geodetic application of inertial navigation, whose performance is impaired as a result of the still somewhat nebulous geodetic models. But naturally any step towards application depends very much on the feed back and cooperation of those who apply, whether in practice or in research. I close with pointing out another aim of the work done in physical geodesy -- and one almost does not dare to mention it in a time of applied and contract research, although it was the fundamental motivation of geodetic work since ancient time, simply to learn more about mother Earth.
Ladies and gentlemen,

Having come to the end of my oration I wish to express my sincere appreciation to her Majesty the Queen of the Netherlands for my appointment as professor of physical geodesy at this University of Technology.

Members of the Executive Board of the University of Technology,

Let me express due appreciation for your confidence in my abilities to live up to the expectations of my function. It is my desire to assure you that I will try to discharge myself of my obligation to the best of my abilities.

Dear colleagues,

I wish to thank for the cordial reception. It is my intent to strive toward a cooperation with you in any manner that could be fruitful to our university and society.

To you professor Ligterink I want to express my special appreciation for the way in which you as Chairman of the Department acted to facilitate the integration.

Professor Baarda,

For your continuing advice and many stimulating discussions, I wish to express my appreciation. They represent a major source of inspiration for my work. I hope to be allowed to rely on them also in the future.

Dear Gerrit Bakker, Professor Koolman, and Guwert Stroo van Haeu,

That your background and interests ideally complement my own, to form so-to-say an alliance of differential geometry and physical geodesy, may be seen as a lucky coincidence. That I could enjoy your support from the very first moment and meanwhile your friendship is certainly more than just a coincidence. To me it is still a great privilege.

To the members of the Department of Geodesy and especially to the members of the "Vakgroep Mathematische en Fysische Geodesy",

I wish to express my appreciation for the excellent cooperation and the intensive exchange of ideas. I am looking forward to a continuation of the same for many years to come.

Ladies and Gentlemen Students,

One of my primary reasons to come from a research institute to the university was to get into contact with students. It is my intent to contribute to the best of my abilities to your professional education in an atmosphere of mutual confidence. In turn I hope to benefit from your critical response. It is this duality of teaching based upon actual research and the critical response to it that seems to me one of the most essential features of university life. That both elements have to possess a dimension of responsibility for society is a precondition especially in difficult times.